

**CREDIT-BASED RECEIVER USING SELECTED TRANSMIT RATES AND STORAGE  
THRESHOLDS FOR PREVENTING UNDER FLOW AND OVER FLOW – METHODS,  
APPARATUS AND PROGRAM PRODUCTS**

5 Inventors: Brian M. Bass, Clark D. Jeffries, and Michael S. Siegel

**BACKGROUND OF INVENTION**

1. Field of Invention:

10 This invention relates to digital communications networks and related apparatus. More particularly, the invention relates to receivers using thresholds on queues of received information to generate transmit rates and communicating those rates to senders for prevention of underflow and overflow of said queues.

15 2. Description of Prior Art:

In Fibre Channel networks, one flow control mechanism for congestion control is the monitoring in a receiver of queue lengths of data awaiting processing by the receiver. The receiver informs the sender of a “credit” value that can be translated into a rate at which the sender sends data. A tutorial on credit based flow control systems for Fibre Channel is described in notes by InterOperability Lab of the University of New Hampshire at

20 [http://www.iol.unh.edu/training/fc/fc\\_tutorial.html#Flow\\_Control](http://www.iol.unh.edu/training/fc/fc_tutorial.html#Flow_Control)

updated 4 May 1998. Typically, the credit or, as it is called herein, transmit rate is determined heuristically and performance may not be guaranteed. What is needed in the art is a receiver which periodically refreshes the transmit rates it generates and thereby provides transmit rates which can be guaranteed to prevent overflow or underflow of the receiver.

5

USP 5,515,359, of Zheng issued May 7, 1996 describes a system for controlling traffic in a digital communication network to avoid data loss due to congestion utilizes an integrated credit-based and rate-based traffic control approach, and adjusts the rate at which data is transmitted from a source in accordance with feedback in the form of rate and credit adjustment information from a network reflecting the ability of the network to transmit data and the destination to receive data. In one embodiment, a source end system sends out resource management cells composed of various fields containing rate and credit information. The intermediate systems of the network and the destination end system update the corresponding fields according to their congestion status and send the resource management cells back to the source end system so that the source end system controls the data transmission accordingly. In a preferred embodiment, a source end system calculates an allowed sending rate from each of the fields in a resource management cell, and the minimum one among them is used to control data transmission.

20

USP 5,777,987 of Adams et al. issued July 7, 1998, describes a method and apparatus for using a primary FIFO and one or more secondary FIFOs in parallel to simplify flow control and routing in packet communication operations wherein at least one FIFO (buffer) is associated with each of a plurality of receiving nodes or components within a receiving node. The received

packets are applied simultaneously to a primary FIFO and to all associated secondary FIFOs in the receiver of a packet communications link. After receipt of a packet, the packet is removed from any secondary FIFOs which correspond to receiver nodes or components to which the packet was not routed. For all receiving nodes or components to which the packet was routed, if

5 the packet was stored in each associated secondary FIFO without overflow, then the packet is also purged from the primary FIFO. If any secondary FIFO overflowed by storage of the received packet, then the packet is purged from the overflowed FIFO and the packet remains stored in the primary FIFO for further processing. Flow control signals are generated and applied to the transmitting source as required in accordance with the status of the primary FIFO.

10 The secondary FIFOs are not directly relevant to flow control logic. The receiving component corresponding to each secondary FIFO locates the next packet for processing by inspecting the associated secondary FIFO as well as the primary FIFO if the secondary FIFO overflowed. These methods and apparatus simplify flow control and routing control in packetized communication receivers

15  
20  
USP 5,748,613 of Kilk, et al. issued May 5, 1998 describes a method of pacing a stream of data transmitted from a data source to a buffered data destination with a determined number of available storage units, the data destinations being configured to consume data and thereby to free storage units for receipt of additional data. The pacing of data communication includes: (1) identifying a beginning transmit rate; (2) incrementing the beginning transmit rate with each storage unit freed to identify an present transmit rate; (3) transmitting units of data in accordance with determined limits, the number of data units sent providing a transmission count; (4) selectively updating the determined number of available storage units by determining the

difference between the beginning transmit rate and the present transmit rate, and determining the sum of the result and the previously determined number of available storage units to provide an updated determined number of available storage units; and (5) selectively updating the determined number of available storage units by determining the difference between the transmission count and the previously determined number of available storage units to provide an updated determined number of available storage units.

USP 6,097,705 of Ben-Michael, et al. issued August 1, 2000, describes a repeater device for forwarding a data packet from a first Ethernet collision domain to a second Ethernet collision domain, the device having a plurality of ports, each port for connection to an independent Ethernet collision domain. Furthermore, each port has an associated receive buffer and an associated transmit buffer, and there is a means for forwarding a data packet from the receive buffer of a receiving port to the transmit buffer of a transmitting port. A data packet received at the receiving port is then first stored in that port's the receive buffer, is forwarded to the transmit buffer of the transmitting port, and is then transmitted from the transmit buffer by the transmitting port.

None of the above prior art discloses a credit based receiver which periodically adjusts transmit rates and storage thresholds and guarantees the prevention of overflow and underflow in the receiver

### INVENTION SUMMARY

A credit based digital communication network is adapted to prevent overflow or underflow of a data storage queue in a receiver by generating a transmit rate value as a feedback

to the sender. The rate adjustments are performed periodically with a fixed time period denoted by  $Dt$ . It is assumed that the sender always has a superabundance of data to send. The data transmit rates are fractions of the maximum possible data transmit speed, designated  $Max$ . The value of  $Dt$  is assumed to be greater than the transmit delay. Since the sender always has a

5 superabundance of data to send, a flow control mechanism can specify any transmit rate up to and including  $Max$  at any time. In a preferred embodiment, the transmit rates are explicitly 0,  $Max/2$ , and  $Max$ . The receiver queue is itself drained at a rate  $R$  that at any time satisfies  $0 \leq R \leq Max$ . The level of occupancy (in bits or other data units) of the receiver storage queue is denoted by  $Q$ . The maximum capacity of the receiving queue is designated  $Q_{max}$ , so at any

10 time,  $0 \leq Q \leq Q_{max}$ . Two thresholds  $T1$  and  $T2$  (with  $0 < T1 < T2 < Q_{max}$ ) of levels of the receiver queue value  $Q$  are determined based upon queuing analysis. A transmit rate is then selected from the possible values by comparison of the receiver queue  $Q$  to the thresholds. The transmit rate value so calculated achieves the desired goals of avoiding overflow and, once the lower threshold has been a positive value at least once, avoiding underflow. As a result the receiver queue is never completely full and never completely empty, regardless of the rate  $R$  at which the receiver queue is drained, with  $0 \leq R \leq Max$ .

### DESCRIPTION OF THE DRAWINGS

The invention will be further understood from the following description of a preferred

20 embodiment taken in conjunction with an appended drawing, in which:

Figure 1 is representation of a credit-based digital communication network adapted to prevent overflow or underflow of a data storage queue in a receiver and incorporating the principles of the present invention.

5 Figure 2 is a flow diagram of a process implemented in the network of Figure 1 for preventing overflow and underflow of a credit-based receiver.

Figure 3A - C are graphs of the network of Figure 1 demonstrating the credit based receiver storage ( $Q_{max}$ ) is not subject to overflow or underflow for different transmit rates and thresholds.

### DESCRIPTION OF PREFERRED EMBODIMENT

In Figure 1, one embodiment of a credit based communication system 100 is disclosed.

10 The system 100 includes a sending station 102 including a transmitter 104 responsive to data packets 106 and a speed or transmit rate controller 108. The speed controller sets the transmit rate ( $Tr$ ) of the transmitter 104 at transmit rates  $Tr = 0$ ,  $Tr = Max/2$  and  $Tr = Max$  where Max denotes the maximum transmit rate of the transmitter. The speed controller is responsive to an updated transmit rate 110 received from a destination 112 via a signal receiving unit 114.

15 The destination 112 receives the data packets 106 in a packet processor 122. The data packets 106 are provided to a buffer queue 124 having a maximum capacity of  $Q_{max}$ . A threshold circuit compares the occupancy of the queue 124 with thresholds  $T1$  and  $T2$  specified by the present invention. The threshold circuit 130 periodically provides an updated transmit rate to a transmit rate register 119. The threshold circuit 130 calculates and sets a lower

threshold T1 and an upper threshold T2 in the buffer 124 based upon Max value 132, Qmax value 134. The details of calculating the thresholds T1 and T2 are based on queuing analysis which will be provided hereinafter. The thresholds T1 and T2 are the thresholds in the buffer 124 used to prevent underflow and overflow, respectively. While two thresholds are described, any number of thresholds may be calculated for the buffer as will be described hereinafter. The transmit rate stored in the register 119 is then periodically communicated from a transmit rate unit 136 to the sender 102 via a communications link 142.

The threshold values T1 and T2 are calculated by a threshold circuit 134 and then compared by the threshold circuit to the level of data packets awaiting processing and temporarily stored in the buffer 124. In one embodiment, if the data packet level stored in the buffer is greater than T2, then an updated transmit rate of 0 is communicated to the sender. Else, if the level of data packets stored in the buffer is greater than T1, then a transmit rate of  $\text{Max}/2$  is communicated to the sender. Else, a transmit rate of Max is communicated to the sender. The transmit rate is processed by a transmit rate unit 136 and communicated over a communications link 142 to a signal receiving unit 114. The flow control transmit rate is updated periodically every  $Dt$  time units and communicated to the sending unit whereupon the sender 102 sends some data at a rate equal to a fraction of the maximum rate to the receiver 112.

Queuing analysis demonstrates underflow and overflow can be prevented in the receiver provided the following conditions are met:

- (1)  $Dt$  is much larger than the time delay of communicating the computed transmit

rate from receiver to sender or time delay in transmission of data from sender to receiver. This is a lower limit on  $Dt$ .

(2) The maximum possible change in one time period  $Dt$  of signal update in receiver queue level is by definition  $Dt \cdot \text{Max}$ . This value should fulfill the inequality  $Dt \cdot \text{Max} < Q_{\text{max}}/8$ .

5 This is an upper limit on  $Dt$ .



The transmit rate signals are as follows:

1. if  $Q \geq T_2$ , then transmit rate = 0
2. else if  $Q \geq T_1$ , then transmit rate =  $\text{Max}/2$
3. else transmit rate = Max.

That is, if  $T_2 \leq Q \leq Q_{\text{max}}$ , then the sender is signaled to send nothing. If  $T_1 \leq Q < T_2$ , then the sender is signaled to send at the rate  $\text{Max}/2$ . If  $0 \leq Q < T_1$ , then the sender is signaled to send at the rate Max.

The present invention also includes specification of the values of  $T_1$  and  $T_2$ . That is,  $T_1$  and  $T_2$  must fulfill conditions (a), (b), and (c), where:

$$T_1 > Q_{\text{max}}/8 \quad \text{Eq(a)}$$

$$T_2 < 15 \cdot Q_{\text{max}}/16 \quad \text{Eq(b)}$$

$$T_1 \leq T_2 - Q_{\text{max}}/16 \quad \text{Eq(c)}$$

**Theorem 1.** The conditions (a), (b), (c) on  $T_1$  and  $T_2$  imply the queue occupancy  $Q$  will never reach  $Q_{\text{max}}$ . Also, once queue occupancy  $Q$  has exceeded 0, it will always thereafter be positive.

Proof:

The maximum value of queue occupancy will occur after a flow control interval  $Dt$  in which the transmit rate was either  $Max/2$  or  $Max$ .

5

Suppose the first case, that is, that maximum value of queue occupancy occurs after a flow control interval in which the transmit rate was  $Max/2$ . Therefore the previous value of  $Q$  was less than  $T2$ . Therefore there must be a positive value  $X$  so that the previous  $Q$  value was  $T2 - X$ . Thus, given a receiver drain rate  $R$  of at least 0, condition (2) on  $Dt$  and condition (b) on  $T2$ , the new  $Q$  value is at most:

$$T2 - X + Max * Dt / 2 - 0 * Dt < T2 - X + Qmax / 16 < 15 * Qmax / 16 + Qmax / 16 = Qmax \quad Eq (1)$$

Suppose the alternative case, namely, that maximum value of queue occupancy occurs after a flow control interval in which the transmit rate was  $Max$ . Then there must be a positive value  $X$  such that the old  $Q$  value is  $T1 - X$ . Given a receiver drain rate  $R$  of at least 0, condition (2) on  $Dt$ , and conditions (b) and (c), the new queue occupancy  $Q$  is at most:

$$T1 - X + Max * Dt - 0 * Dt < T2 - Qmax / 16 - X + Qmax / 8 < Qmax. \quad Eq (2)$$

20

Thus, in both cases, the deductions of equations 1 and 2 show the maximum queue occupancy possible ever is less than  $Qmax$ .

Concerning underflow, after  $Q$  occupancy has been positive at least once, a transmit rate of  $Max$  could not lead to a decrease in  $Q$ , given the restrictions on  $R$ . Therefore the minimum  $Q$  occupancy will occur after a flow control interval  $Dt$  in which the transmit rate is either 0 or  $Max/2$ .

5

Suppose the first case, that is, that minimum value of queue occupancy occurs after a flow control interval in which the transmit rate was 0. Therefore there must be a nonnegative value  $X$  so that the old  $Q$  value was  $T2+X$ . Since the drain rate  $R$  from the receiver is at most  $Max$ , the new  $Q$  value is at least  $T2 + X - Max \cdot Dt$ . Given condition (2) on  $Dt$  and conditions (a) and (b), the new  $Q$  value is at least:

$$T2 + X - Q_{max}/8 \geq T1 + Q_{max}/16 - Q_{max}/8 > Q_{max}/8 - Q_{max}/16. \quad \text{Eq (3)}$$

Therefore the new  $Q$  value is at least  $Q_{max}/16$ .

Suppose the alternative case, namely, that minimum value of queue occupancy occurs after a flow control interval in which the transmit rate was  $Max/2$ . Therefore there must be a nonnegative value  $X$  so that the old  $Q$  value was  $T1 + X$ . With a drain rate  $R$  limited by  $Max$ , the new  $Q$  value is at least  $T1 + X - Max \cdot Dt$ . Given condition (2) on  $Dt$  and condition (a), the new  $Q$  value is therefore greater than  $Q_{max}/8 - Q_{max}/8 = 0$ . Therefore the new  $Q$  value is greater than 0. This deduction and the deduction leading to equation 3 show that the new  $Q$  value must be positive.

End of proof of theorem.

Alternative embodiment: More thresholds could be specified with a finer granularity of transmit rates. For example, let  $N$  positive thresholds  $0 < T_1 < T_2 < \dots < T_N < Q_{\max}$  satisfy

$$Q_{\max} * (1/16 + i/16) < T_i < Q_{\max} * (1 - 1/(8*i)) \text{ for } i = 1, 2, \dots, N \quad \text{Eq (4)}$$

5

Let flow rates be specified as functions of  $Q$  as follows:

1. if  $Q \geq T_N$  then transmit rate = 0
2. else if  $Q \geq T_{N-1}$  then transmit rate =  $\text{Max}/N$
3. else if  $Q \geq T_{N-2}$  then transmit rate =  $\text{Max}/(N-1)$
4. else if  $Q \geq T_{N-3}$  then transmit rate =  $\text{Max}/(N-2)$
5. else if  $Q \geq T_i$  then transmit rate =  $\text{Max}/(i+1)$
6. else if  $Q \geq T_2$  then transmit rate =  $\text{Max}/3$
7. else if  $Q \geq T_1$  then transmit rate =  $\text{Max}/2$
8. else transmit rate =  $\text{Max}$

Theorem 2. Given the above N thresholds and rates, the queue occupancy Q will never reach Qmax. Also, once queue occupancy has exceeded 0, it will always thereafter be positive.

Proof.

5

Suppose previous Q was at or above TN. Then the rate was 0, so no increase is possible.

Suppose previous Q was just below Ti, so rate was Max/i. Then the maximum value Q can attain less than

$$T_i + \text{Max} * \text{Dt} / i < Q_{\text{max}} * (1 - 1/(8*i)) + (\text{Max}/i) * (Q_{\text{max}}/(8*\text{Max})) = Q_{\text{max}}. \quad \text{Eq (5)}$$

10 Thus, the deduction leading to equation 5 shows overflow is impossible.

Suppose the previous Q was at or above Ti, so rate is Max/(i+1). Then the smallest Q can be is

$$T_i + \text{Max} * \text{Dt} / (i+1) - \text{Max} * \text{Dt} > Q_{\text{max}} * (1/16 + i/16) + (Q_{\text{max}}/8) * (1/(i+1)) - Q_{\text{max}}/8 = Q_{\text{max}} * (-1/16 + 1/(8*(i+1)) + i/16) > 0. \quad \text{Eq (6)}$$

15 Suppose previous Q was below T1. Then the rate is Max, which is greater than or equal to the drain rate R, so depletion to zero is impossible. Thus, this observation and the deduction leading to equation 6 show underflow is impossible.

End of proof of theorem.

20

Now turning to Figure 2 a process 200 will be described in conjunction with Figure 1 for implementing the prevention of underflow and overflow in a credit-based receiver in the digital communication network 100 of Figure 1, as follows:

In step 202, data packets are transmitted to a destination 112. The rate of transmission is controlled by a speed controller 108.

5 In step 204, the data packets are received by a packet processor 122 and temporarily stored in buffer 124 while awaiting further processing. The updated transmit rate is determined by comparing buffer queue level  $Q$  with maximum capacity  $Q_{\max}$  132 and with thresholds  $T_1$  and  $T_2$  in the threshold circuit 130, and then the updated transmit rate is stored in register 119 in step 206.

10 In step 208 the updated transmit rate is communicated from the receiver 112 by action of the transmit rate unit 136 to the signal receiving unit 114 in the sender 102.

15 In step 210 the updated transmit rate 110 is stored in a register. In step 212 a speed controller 108 transmits data at the updated transmit rate from the transmitter 104.

Thresholds  $T_1$  and  $T_2$  are computed at initialization time from  $Q_{\max}$  in a threshold circuit 132 by means of three equations described herein, namely,

20  $T_1 > Q_{\max}/8$  Eq(a)

$T_2 < 15 \cdot Q_{\max}/16$  Eq(b)

$$T1 \leq T2 - Q_{\max}/16$$

Eq(c)

5        Figures 3A-C show test results of a credit-based receiver implementing the principles of the present invention and demonstrating overflow and underflow are prevented in the receiver.

Figure 3A shows time units along the horizontal axis and the queue processing or service rate R of a queue on the vertical axis. In the example, the value of  $Q_{\max}$  is 1 unit of data (for example, one megabit). The maximum service rate Max of the queue is .125 data units per time step  $Dt$ . The variable service rate R 302 is shown. For 500 time steps of duration  $Dt$ , the rate R is random and between 0 and .04 in units of data per time. For the next 500 time steps the rate R is constantly equal to the maximum rate  $Max = .125$  data units per time step.

Figure 3B shows the transmit rates 304 calculated by the algorithm for the above conditions. Note that the transmit rate is 0 or  $Max/2$  for the first 500 time steps. Then, as the service rate changes, the transmit rate is  $Max/2$  or  $Max$  for the next 500 time steps.

Figure 3C shows the level of queue occupancy 106 during the above experiment. Not that the value of Q does not enter into an overflow state. Likewise the value of Q does not enter into an underflow condition.

Thus, Figures 3A-3C demonstrate that the setting of transmit rate in accordance with the capacity of the buffer queue 124 relative to the thresholds  $T1$  and  $T2$  prevents the buffer from underflowing or overflowing for the transmit fractions.

While the invention has been described in conjunction with a preferred embodiment various changes can be made without departing from the spirit and scope of the invention as defined in the appended claims, in which:

We claim: